

University of Dundee

Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes

Trinkle-Mulcahy, Laura; Boulon, Severine; Lam, Yun Wah; Urcia, Roby; Boisvert, Francois-Michel ; Vandermoere, Franck

Published in:
Journal of Cell Biology

DOI:
[10.1083/jcb.200805092](https://doi.org/10.1083/jcb.200805092)

Publication date:
2008

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in Discovery Research Portal](#)

Citation for published version (APA):

Trinkle-Mulcahy, L., Boulon, S., Lam, Y. W., Urcia, R., Boisvert, F-M., Vandermoere, F., Morrice, N. A., Swift, S., Rothbauer, U., Leonhardt, H., & Lamond, A. (2008). Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *Journal of Cell Biology*, 183(2), 223-239.
<https://doi.org/10.1083/jcb.200805092>

General rights

Copyright and moral rights for the publications made accessible in Discovery Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from Discovery Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes

Laura Trinkle-Mulcahy,^{1,3} Séverine Boulon,¹ Yun Wah Lam,⁴ Roby Urcia,⁵ François-Michel Boisvert,¹ Franck Vandermoere,² Nick A. Morrice,² Sam Swift,¹ Ulrich Rothbauer,⁶ Heinrich Leonhardt,⁶ and Angus Lamond¹

¹Wellcome Trust Centre for Gene Regulation and Expression, and ²MRC Protein Phosphorylation Unit, University of Dundee, Dundee DD1 5EH, Scotland, UK

³Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, Canada K1H 8M5

⁴Department of Biology and Chemistry, City University of Hong Kong, Hong Kong

⁵Department of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, Scotland, UK

⁶Department of Biology, LMU Biozentrum, 82152 Planegg-Martinsried, Germany

The identification of interaction partners in protein complexes is a major goal in cell biology. Here we present a reliable affinity purification strategy to identify specific interactors that combines quantitative SILAC-based mass spectrometry with characterization of common contaminants binding to affinity matrices (bead proteomes). This strategy can be applied to affinity purification of either tagged fusion protein complexes or endogenous protein complexes, illustrated here using the well-characterized SMN complex as a model. GFP is used as the tag of choice because it shows minimal non-

specific binding to mammalian cell proteins, can be quantitatively depleted from cell extracts, and allows the integration of biochemical protein interaction data with in vivo measurements using fluorescence microscopy. Proteins binding nonspecifically to the most commonly used affinity matrices were determined using quantitative mass spectrometry, revealing important differences that affect experimental design. These data provide a specificity filter to distinguish specific protein binding partners in both quantitative and nonquantitative pull-down and immunoprecipitation experiments.

Introduction

Most biological processes involve the action and regulation of multiprotein complexes. In many cases, separate properties such as subcellular localization, catalytic activity, and substrate specificity are determined by different polypeptides in a holoenzyme complex, and specific protein interaction partners may be present in nonstoichiometric amounts. For example, catalytic subunits such as protein phosphatase 1 (PP1) can interact with a spectrum of alternative protein partners, which thus bind nonstoichiometrically to generate a range of holoenzymes with different specificities (for review see Moorhead et al., 2007). This can make it difficult to distinguish specific but low abundance interacting proteins from the larger number of low affinity, but abundant, contaminant proteins that are inevitably recovered using commonly used methods such as pull-down or immunoprecipitation strategies. A key goal in most areas of cell biology, therefore, is the characterization of the protein components of

multiprotein complexes through the reliable identification of specific protein interaction partners.

Any putative interaction partner identified either through affinity purification or biochemical fractionation must be validated to confirm its physiological relevance. These downstream validation experiments, involving detailed molecular characterization, are both costly and time consuming and thus it is imperative to focus resources on those subsets of potential interactions with a high probability of biological significance. Continuing improvement in the sensitivity and resolution of the mass spectrometric technology for protein identification, for example, allows for the identification of ever larger numbers of proteins in immunoaffinity and pull-down experiments. In addition to bona fide interaction partners, however, these expanding lists include increased numbers of contaminant proteins, including those that bind nonspecifically to the affinity matrix. The problem of nonspecific binding cannot be overcome satisfactorily

Correspondence to Laura Trinkle-Mulcahy: ltrinkle@uottawa.ca; or Angus Lamond: angus@lifesci.dundee.ac.uk

Abbreviations used in this paper: FP, fluorescent protein; SILAC, stable isotope labeling with amino acids in cell culture.

The online version of this article contains supplemental material.

© 2008 Trinkle-Mulcahy et al. This article is distributed under the terms of an Attribution-Noncommercial-Share Alike-No Mirror Sites license for the first six months after the publication date (see <http://www.jcb.org/misc/terms.shtml>). After six months it is available under a Creative Commons License (Attribution-Noncommercial-Share Alike 3.0 Unported license, as described at <http://creativecommons.org/licenses/by-nc-sa/3.0/>).

Supplemental Material can be found at:
<http://jcb.rupress.org/content/suppl/2008/10/20/jcb.200805092.DC1.html>

using high stringency purification methods; although this can reduce the level of nonspecific binding, it will inevitably also remove low abundance and low affinity specific partner proteins. The most effective strategy must therefore preserve all specific interaction events, which inevitably results in a large number of nonspecific proteins also copurifying that must be identified and discarded.

To solve this problem, we and others have demonstrated that a quantitative mass spectrometry-based approach combined with isotope labeling can help to distinguish which of the many proteins identified in a pull-down or immunoprecipitation experiment represent specific binding. This is done by the inclusion of a negative control, which provides a background of contaminant proteins that bind nonspecifically to the affinity matrix and/or the fusion tag, against which proteins that bind specifically to the protein of interest clearly stand out (for review see Vermeulen et al., 2008). For example, using a combination of stable isotope labeling with amino acids in cell culture (SILAC)-based quantitative proteomics (Ong et al., 2002) with immunoprecipitation of GFP-tagged fusion proteins, we revealed differences in binding partners for two different isoforms of the nuclear protein phosphatase, PP1 (Trinkle-Mulcahy et al., 2006). Other groups have used a similar approach based on tagged bait proteins to map the spectrum of human 26S proteasome interacting proteins (Wang and Huang, 2008) and to detect dynamic members of transcription factor complexes (Mousson et al., 2008). Isotope-based quantitative approaches have also been used to define tagged protein complexes in yeast (Ranish et al., 2003; Tackett et al., 2005) and both tagged and endogenous protein complexes in mammalian cells (Blagoev et al., 2003; Cristea et al., 2005; Selbach and Mann, 2006).

Although the isotope labeling strategy used in a SILAC affinity purification approach provides great help in separating specific from nonspecific interactors, experience shows that not all specific interactions can be unambiguously determined, particularly near the threshold level where signal-to-noise ratios are close to background. Here we describe a new SILAC-based mass spectrometry strategy that specifically addresses this issue, incorporating methods to increase the signal, i.e., the abundance of purified protein complexes, while reducing or filtering out the noise, i.e., proteins that bind nonspecifically to the affinity matrix, tag, and/or antibody.

The efficiency of detecting interaction partners relies upon efficient depletion of the targeted complex. Here we show that GFP-tagged proteins can be near quantitatively depleted using the recently developed GFP binder (Rothbauer et al., 2008). The GFP binder is an *Escherichia coli*-expressed 16-kD protein derived from a llama heavy chain antibody that binds with high affinity and specificity to GFP. This underlines the utility of using GFP as a dual tag for both affinity purification and in vivo fluorescence microscopy. Furthermore, characterizing the proteins that bind nonspecifically to three of the most commonly used affinity matrices, in either whole cell, nuclear, or cytoplasmic extracts of mammalian cells, provides a “bead proteome” filter. This facilitates distinguishing specific from nonspecific binding proteins and thereby allows objective prioritization of suitable targets for detailed molecular characterization.

In summary, we present here a powerful and reliable workflow that can be applied to analyze affinity-purified protein complexes isolated using either tagged fusion proteins or via immunoprecipitation of endogenous proteins.

Results

Optimized workflow for quantitative analysis of endogenous and tagged protein complexes

A standard workflow for SILAC-based analysis of protein interaction partners in pull-down experiments is summarized in Fig. 1. In brief, the total protein components isolated from either an immunoprecipitation or affinity pull-down experiment are size fractionated using SDS-PAGE. The gel is cut into typically 5–10 slices, each of which is digested with trypsin and the resulting peptides eluted and analyzed by high sensitivity mass spectrometry (see Materials and methods).

The procedures described show the optimized protocols we have derived from over 50 separate interaction analyses. This is applied routinely for the analysis of interaction partners binding to fluorescent protein (FP)-tagged fusion proteins in whole cell, cytoplasmic, and nuclear extracts (Fig. 1 A). Cells expressing the tagged protein are grown in “heavy” media, i.e., containing ^{13}C -substituted arginine and lysine. As a control, either parental/untransfected cells or cells expressing free GFP are grown in “light”, i.e., unlabeled (^{12}C) media. Initially, cell lines expressing free GFP were routinely used as a control. However, experience showed that the level of nonspecific protein binding to free GFP in mammalian cell lines was so low that nonexpressing cells can also provide a suitable negative control.

In this approach the negative “light” control and the experimental “heavy” sample are mixed before mass spectrometric analysis. This reduces the effective experimental variability that inevitably results when the samples are processed independently. Here extracts mixed before the GFP immunoprecipitation step were analyzed. However, separate immunoprecipitations can also be performed and the affinity matrices mixed before eluting proteins for further analysis. Specific steps in the protocol can be optimized according to the specific requirements of individual experiments. However, it is recommended that the duration of incubation for the binding step to the affinity matrix is always minimized, to reduce potential losses of dynamic or weakly associated factors. The present protocol has been optimized using extracts from HeLa and U2OS cells. Analysis using extracts from other cell lines should be optimized individually to ensure efficient protein recovery.

A similar SILAC strategy can also be applied for the analysis of protein interaction partners recovered from direct immunoprecipitation of endogenous complexes (Fig. 1 B). In this case, a control must be performed with a nonspecific antibody, e.g., either preimmune IgG, or an antibody raised against a tag or epitope that is not expressed in these cells. Because separate, parallel immunoprecipitations are required for the control and test samples, care must be taken when mixing the beads to ensure that equal quantities of material are compared.

An important issue for maximizing the identification of protein interaction partners is ensuring both efficient isolation of

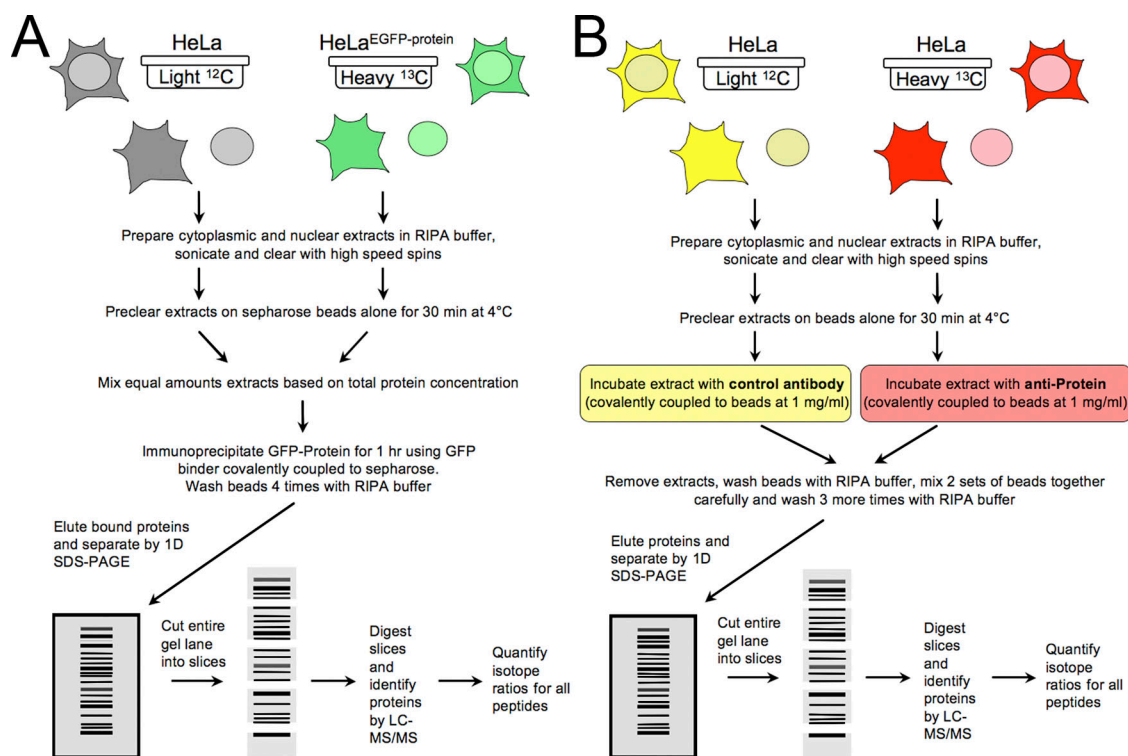


Figure 1. Protocols used for SILAC-based analysis of protein interaction partners in pull-down experiments. (A) HeLa cells expressing a GFP-tagged protein are metabolically labeled by culturing in “heavy” media containing ^{13}C -isotopes of arginine and lysine, while the parental HeLa cells are grown in “light” media containing the ^{12}C -isotopes of arginine and lysine. Whole cell extracts can be prepared or, as shown here, cells can be fractionated for preparation of separate cytoplasmic and nuclear extracts. In this case, extracts are pre-cleared on Sepharose beads and then mixed in equal amounts before affinity purification of the GFP-tagged protein using the GFP binder (1 h incubation). Proteins are eluted from the beads and separated by 1D SDS-PAGE for digestion and LC-MS/MS analysis. (B) For SILAC analysis of an endogenous protein, two populations of HeLa cells are grown in light and heavy media, respectively, before harvesting and preparation of cellular extracts. Equal total protein amounts of each extract are subjected to separate immunoprecipitation experiments, either using an antibody to the protein of interest or a control antibody covalently bound to beads at an equivalent concentration. The separate immunoprecipitates are mixed carefully to minimize variability and the proteins eluted and analyzed as described above.

the target protein under study and achieving a high signal-to-noise ratio. In the case of FP-tagged proteins, our results show this is best achieved using the recently developed GFP binder (Rothbauer et al., 2008), which reproducibly provides near-quantitative depletion of GFP fusion proteins (Fig. 2). Direct comparison with commercially available anti-GFP monoclonal antibodies (mAbs) shows that an affinity matrix coupled to the GFP binder routinely produces higher depletion efficiencies and improves signal-to-noise ratios (Fig. 2, A and B; and unpublished data).

GFP is a 27-kD protein, and a tag of this size could potentially bind itself to a range of cell proteins. We note that in vivo FRAP measurements in both the cytoplasm and nucleus show that photobleaching GFP expressed in live cells results in rapid recovery (Fig. S1, available at <http://www.jcb.org/cgi/content/full/jcb.200805092/DC1>). This indicates that GFP in vivo predominantly diffuses as a free protein and therefore binds weakly or not at all with most cellular protein complexes. Nonetheless, a subset of GFP molecules could still associate with cell proteins, and it is also possible that this could increase upon cell fractionation. To test this more rigorously, the SILAC pull-down method was used to analyze directly which proteins in mammalian cell extracts copurify with GFP isolated using either the GFP binder or a commercially available anti-GFP mAb (Fig. 2 C). Data from four independent experiments generated a short list

of potential GFP-interacting proteins that should be considered as possible contaminants when identified in any interaction analysis of a GFP-tagged protein. However, none of these putative contaminants were recovered in all four experiments and most are also identified as proteins that bind nonspecifically to affinity matrices (see below). Consistent with the FRAP data, it was observed in the extracts tested that there are no major contaminating proteins that copurify reproducibly with free GFP. However, attention is drawn to six proteins, specifically variants of heat shock 70-kD protein, cytokeratins 8 and 18, and ubiquitin, which were most frequently detected as copurifying with GFP-tagged fusion proteins (Fig. 2 C). It is possible that these proteins, which all bind nonspecifically to the Sepharose matrix, are not binding GFP directly but are instead up-regulated in the cell line overexpressing GFP. In summary, the SILAC data demonstrate that GFP, despite its size of 27 kD, is an effective tag for use in pull-down experiments. It shows low levels of non-specific interactions and can be quantitatively depleted from cell extracts using the GFP binder.

Characterization of Sepharose bead proteome

Next, a systematic assessment was made of which proteins in cell extracts bind nonspecifically to the Sepharose matrix, which

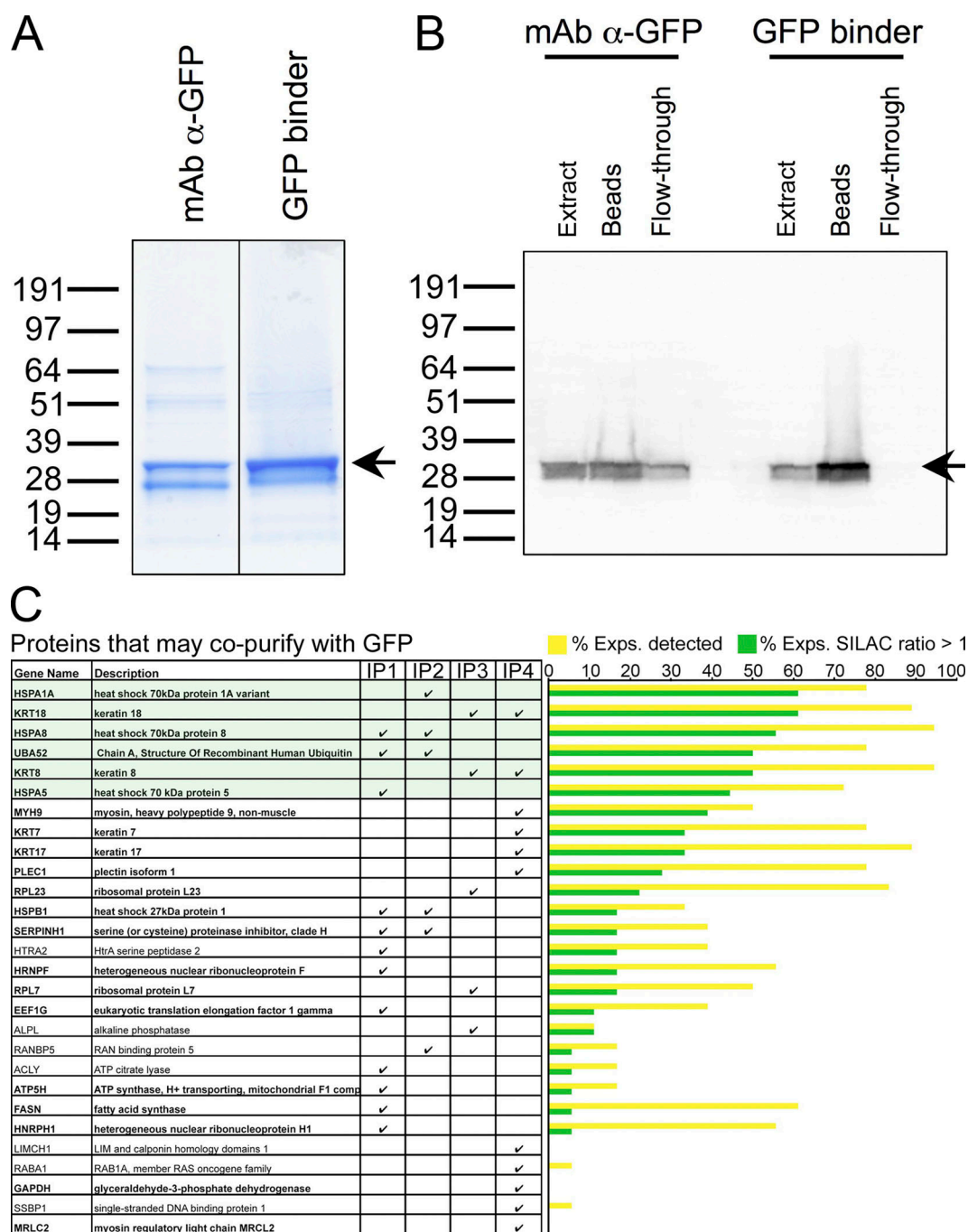


Figure 2. GFP as a tag in immunoaffinity experiments. Although a commercial monoclonal anti-GFP antibody is capable of isolating significant amounts of free GFP from a stable HeLa cell line, the GFP binder is more efficient, as demonstrated both by Coomassie staining of protein eluted from the affinity matrices (A) and Western blotting using anti-GFP antibodies (B). Whether the mAb or GFP binder is used to purify GFP, there are very few proteins that bind nonspecifically to this tag (C). Four independent experiments were performed to identify proteins that may copurify with GFP, as indicated by SILAC ratios greater than 1 (IP1: whole cell extract, GFP binder; IP2: whole cell extract, monoclonal anti-GFP antibody; IP3: cytoplasmic extract, monoclonal anti-GFP antibody; IP4: nuclear extract, monoclonal anti-GFP antibody). No one protein was identified in every experiment, and most of them (in bold) have been identified as binding nonspecifically to the Sepharose bead matrix. This list was then screened against a set of 18 independent GFP protein immunoaffinity experiments performed using the GFP binder for purification and parental cells as the negative control. Proteins were scored for the percentage of experiments in which they were detected (yellow), and for the percentage of experiments in which they were detected and showed a SILAC ratio greater than 1 (green). Six proteins, representing three protein classes (heat shock/chaperone, cytokeratin, and ubiquitin), have been highlighted in green as the most frequently detected and potentially able to bind GFP.

has been used routinely in pull-down experiments and with the GFP binder (Tables I and II). We define the set of proteins binding to the affinity matrix as a “bead proteome.” Data were pooled from 27 independent SILAC pull-down experiments on 11 separate

GFP fusion proteins in either whole cell, cytoplasmic, or nuclear extracts prepared from HeLa and U2OS cells using standard RIPA buffer (see Materials and methods). Analysis of the combined dataset reveals a wide range of cellular proteins that

Table I. **Sepharose bead proteome: most common protein classes**

Protein class	Most commonly found
Cytoskeletal/structural/ motility proteins	Actin
	Cofilin
	Desmin
	Desmoplakin
	Epiplakin
	Filamin
	Myosin
	Peripherin
	Plectin
	Tropomyosin
	Tubulin
	Vimentin
DEAD box proteins	
Eukaryotic translation elongation and initiation factors	
Heat shock proteins	
Histones	
hnRNP proteins	
Ribosomal proteins	

Proteins that bind nonspecifically to Sepharose fall into several distinct classes, as shown here, and are found in nearly every SILAC immunoprecipitation experiment carried out using Sepharose as an affinity matrix.

routinely bind to the Sepharose matrix and which therefore must be regarded as potential nonspecific contaminants whenever they are identified in protein interaction studies. These include histones, hnRNP proteins, heat shock proteins, ribosomal proteins, translation and initiation factors, DEAD box proteins, and multiple cytoskeletal proteins (Table I). Over 100 additional proteins of other classes were also identified (Table II). These common matrix-binding contaminants have therefore been incorporated into a filter set that can be used to compare with sets of proteins identified as potential specific interaction partners for any target protein under study.

Comparison of Sepharose, agarose, and magnetic bead proteomes

Using the SILAC protocol, a comparison was made of nonspecific protein binding to Sepharose as compared with two other commonly used affinity matrices, i.e., agarose and magnetic beads (Fig. 3). In this case, labeling was conducted using three isotopic states, i.e., ^{12}C -arg and ^{12}C -lys for agarose, ^{13}C -arg and D4-lys for Sepharose, and $^{13}\text{C}/^{15}\text{N}$ -arg and $^{13}\text{C}/^{15}\text{N}$ -lys for magnetic beads. Nonspecific protein binding was observed for all three matrices after incubation of either nuclear or cytoplasmic extracts, whether the incubation time was short (30 min) or long (18 h). At both the short and long time points, a similar distribution of classes of contaminating proteins was observed, although the levels of protein binding can increase after longer incubation. An interesting difference was apparent in the relative performance of Sepharose and magnetic beads when incubated with either nuclear or cytoplasmic extracts. Thus, magnetic beads, which showed more nonspecific binding to structural/motility protein classes and lower nonspecific binding to nucleic acid-binding factors, had lower backgrounds of contaminating pro-

teins in nuclear extracts as compared with Sepharose. In contrast, Sepharose, which showed more nonspecific interactions with nucleic acid-binding factors, gave better results than magnetic beads in reducing nonspecific background in cytoplasmic extracts (Fig. 3 C; Table S1, available at <http://www.jcb.org/cgi/content/full/jcb.200805092/DC1>). In the case of agarose beads, similar levels of nonspecific binding to Sepharose were observed in nuclear extracts, whereas agarose beads showed lower nonspecific binding in cytoplasmic extracts as compared with either Sepharose or magnetic beads (Fig. S2, available at <http://www.jcb.org/cgi/content/full/jcb.200805092/DC1>). Overall, it can be concluded that the affinity matrices constitute a major source of nonspecific protein binding for all protein interaction studies and the detailed data obtained from comparing the three main types of affinity matrices show that no single type of bead is ideally suited to all applications. Rather, improved results with respect to nonspecific protein binding can be obtained by using different types of affinity matrix depending upon whether protein interaction studies are performed using cytoplasmic or nuclear extracts, or other types of cellular fractions.

Application of SILAC strategy to identify protein interaction partners

Having identified parameters affecting nonspecific protein binding, the optimized workflow described above was tested for the analysis of a previously characterized multiprotein complex. As a model system, we selected for analysis the intensively studied and well-characterized SMN complex. SMN is the product of the major human gene responsible for the inherited genetic disorder spinal muscular atrophy (for review see Kolb et al., 2007) and is known to form a complex with multiple specific partner proteins, including gemins and snRNP proteins (see Table III and references therein).

Because SMN is found in multiprotein complexes in both the nucleus and the cytoplasm (Fig. 4 A), and because some of its previously identified interactions were reported to be compartment specific (Fig. 4 B), we fractionated cells into nuclear and cytoplasmic extracts to compare the interaction partners identified by SILAC in both compartments. A HeLa cell line stably expressing GFP-SMN (Sleeman et al., 2003) was grown in media containing ^{13}C -labeled arginine and lysine, with parental HeLa cells grown in normal ^{12}C -labeled media as a negative control. The cells were harvested and fractionated into cytoplasmic and nuclear extracts, pull-down experiments were performed using the GFP binder, and proteins were analyzed by mass spectrometry. This resulted in identification of over 20 proteins previously described to copurify with SMN. The average SILAC ratio and number of peptides identified for each protein in both cytoplasmic and nuclear extracts is listed in Table III.

To facilitate identification of specific binding partners, we used a data analysis approach that incorporated both SILAC ratios (i.e., $^{13}\text{C}:^{12}\text{C}$ peptide ratios) and relative peptide abundance (Fig. 4, C and D). These data plotting log SILAC ratios versus total peptide intensity show that SMN itself and the known core members of the SMN protein complex (e.g., gemins 2–8, shown in yellow in Fig. 4, C and D) are readily identified.

These data also show that p80 coilin, which was previously shown to interact with SMN specifically in the nucleus, was here

Table II. **Sepharose bead proteome: other proteins of additional classes**

Gene Name	Description	NP	CP	WC
		% Exp.	% Exp.	% Exp.
ADAR	Double-stranded RNA-specific adenosine deaminase	54.5	0.0	11.1
AHNAK	AHNAK nucleoprotein isoform 1	54.5	28.6	55.6
ALB	Albumin	100.0	45.5	57.1
ANXA1, 2	Annexin 1, A2	100.0	100.0	77.8
ASCC3L1	Activating signal co-integrator 1 complex subunit 3-like 1	63.6	0.0	22.2
ASS1	Argininosuccinate synthase	0.0	57.1	55.6
ATAD	ATPase family, AAA domain containing protein	27.3	0.0	22.2
ATP5A, 5B	ATP synthase, H ⁺ transporting, mitochondrial F1 complex	81.8	14.3	77.8
BAG2	BCL2-associated athanogene 2	27.3	0.0	11.1
BOLA2B	BolA-like protein 2B	22.2	36.4	28.6
CAD	Carbamoylphosphate synthetase 2	0.0	85.7	33.3
CAND1	Cullin-associated NEDD8-dissociated protein 1	9.1	28.6	33.3
CAPRN1	Cytoplasmic activation- and proliferation-associated protein 1	33.3	45.5	0
CCT	Chaperonin containing TCP1	45.5	28.6	55.6
CD180	Elongation factor 1-alpha	45.5	57.1	33.3
CFL1	Cofilin	63.6	85.7	66.7
CHD3, D4	Chromodomain-helicase-DNA-binding protein 3, 4	18.2	0.0	11.1
CLEC2D	C-type lectin domain family 2, member D	36.4	0.0	0.0
CLTC	Cathrin heavy chain 1	36.4	85.7	77.8
COPA, B1	Coatomer protein complex, subunits A, B1	18.2	28.6	44.4
CORO1C	Coronin, actin binding protein 1C	66.7	27.3	14.3
CPS1	Carbamoyl-phosphate synthetase 1	54.5	71.4	44.4
CRKL	Crk-like protein	22.2	9.1	28.6
CSDA	Cold shock domain-containing protein A	55.6	36.4	71.4
CSRP2	Cysteine and glycine-rich protein 2	54.5	28.6	11.1
DBN1	Drebrin 1 (developmentally regulated brain protein)	66.7	36.4	28.6
DHRS2	Dehydrogenase/reductase (SDR family) member 2	36.4	14.3	22.2
DUT	dUTP pyrophosphatase	22.2	18.2	14.3
DYNLL1	Dynein light chain 1	27.3	28.6	11.1
EDARRAD	EDAR-associated death domain	9.1	0.0	66.7
ELAVL1	ELAV-like 1	63.6	0.0	22.2
EMD	Emerin	36.4	28.6	0.0
ENO	Enolase 1	9.1	14.3	77.8
EWSR1	Ewing sarcoma breakpoint region 1	27.3	14.3	33.3
FARSA,B,FASN	Phenylalanyl-tRNA synthetase; Fatty acid synthase	36.4	85.7	88.9
FBL	Fibrillarin	63.6	14.3	22.2
FKSG30	FKSG30	36.4	57.1	66.7
FUS	Fus-like protein	36.4	14.3	33.3
G3BP1,2	GTPase activating protein (SH3 domain) binding protein 1, 2	44.4	18.2	14.3
G6PD	Glucose-6-phosphate 1-dehydrogenase	9.1	28.6	22.2
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	18.2	42.9	55.6
GFAP	Glial fibrillary acidic protein	36.4	42.9	55.6
GNAS	Guanine nucleotide-binding protein G	22.2	9.1	14.3
GNB2L1	Guanine nucleotide-binding protein (G protein), beta polypeptide 2-like 1	36.4	28.6	44.4
GNL2,3	Guanine nucleotide-binding protein-like 2 (nucleolar)	100	0	0
GSR	Glutathione reductase	36.4	42.9	33.3
GSTM3	Glutathione S-transferase M3	18.2	28.6	22.2
ILF 2, 3	Interleukin enhancer binding factor 2, 3	54.5	42.9	55.6
KHDRBS1,2,3	KH domain containing, RNA binding, signal transduction associated	45.5	0.0	11.1
KHSRP	Far upstream element-binding protein 2	36.4	0.0	11.1
KIF2,4	Kinesin family member 2, 4	44.4	36.4	14.3
LDHA	Lactate dehydrogenase A	9.1	42.9	55.6
LGALS1,3	Beta-galactoside-binding lectin	72.7	57.1	55.6
LIMA1	LIM domain and actin-binding protein 1	44.4	9.1	14.3
LMNA, B	Lamin A/C, B	90.9	0.0	33.3
MATR3	Matrin 3	63.6	0.0	44.4

Table II. **Sepharose bead proteome: other proteins of additional classes** (Continued)

Gene Name	Description	NP	CP	WC
		% Exp.	% Exp.	% Exp.
MCM3,5	Minichromosome maintenance complex component 3, 5	54.5	14.3	44.4
MIF	Macrophage migration inhibitory factor	9.1	28.6	22.2
MSH2	MutS protein homologue 2	77.8	0	14.3
MTCH2	Mitochondrial carrier homologue 2	36.4	0.0	11.1
NACA	Nascent polypeptide-associated complex subunit alpha	11.1	18.2	14.3
NCL	Nucleolin	72.7	42.9	22.2
NES	Nestin	55.6	9.1	14.3
NME1, 2	Non-metastatic cells protein 1, 2	18.2	28.6	11.1
NONO	Non-POU domain containing, octamer-binding	36.4	0.0	11.1
NPM1	Nucleophosmin 1	81.8	0.0	44.4
NUDT16L1	Nudix (nucleoside diphosphate linked moiety X)-type motif 16-like 1	77.8	0	14.3
NUMA1	Nuclear mitotic apparatus protein 1	66.7	18.2	42.9
NUP155	Nucleoporin 155 kD	54.5	0.0	11.1
PABPC1,3,4	Poly(A) binding protein, cytoplasmic 1, 3, 4	18.2	57.1	33.3
PALLD	Palladin, cytoskeletal associated protein	45.5	0.0	11.1
PARK7	DJ-1 protein	36.4	28.6	44.4
PARP1	Poly (ADP-ribose) polymerase family, member 1	27.3	0.0	11.1
PCBP1, 2	Poly(rC)-binding protein 1, 2	63.6	85.7	55.6
PCMT	Protein-L-isoaspartate(D-aspartate) O-methyltransferase	18.2	57.1	33.3
PDIA6	Protein disulfide-isomerase A6	63.6	14.3	77.8
PDLIM2,4	PDZ and LIM domain protein 2	55.6	0	0
PFDN2	Prefoldin subunit 2	9.1	14.3	11.1
PFN2	Profilin 2	36.4	28.6	33.3
PHB, PHB2	Prohibitin, prohibitin 2	63.6	42.9	77.8
PHF5A	PHD finger protein 5A	45.5	0.0	11.1
PHGDH	Phosphoglycerate dehydrogenase	18.2	71.4	44.4
PKM2	Pyruvate kinase, muscle	54.5	71.4	66.7
POTE2	Protein expressed in prostate, ovary, testis, and placenta 2	54.5	42.9	66.7
PPIA	Peptidylprolyl isomerase A (cyclophilin A)	66.7	27.3	71.4
PRDX1,2,3,4	Peroxiredoxin 1, 2, 3, 4	90.9	71.4	100.0
PRKDC	Protein kinase, DNA-activated	81.8	14.3	44.4
PTBP1,2	Polypyrimidine tract-binding protein 1, 2	63.6	14.3	11.1
RALY	RNA-binding protein (autoantigenic, hnRNP-associated with lethal, yellow)	100	27.3	28.6
RCC2	Regulator of chromosome condensation 2	63.6	0.0	11.1
S100A6,9,10,14	S100 calcium binding protein A	100	36.4	42.9
SAP18	Sin3-associated polypeptide, 18 kD	44.4	0	14.3
SEC61B	Protein transport protein Sec61 beta subunit	45.5	0.0	11.1
SERBP1	SERPINE1 mRNA binding protein 1	45.5	0.0	22.2
SERPINH1,A11	Serine (or cysteine) proteinase inhibitor	72.7	28.6	66.7
SF3B	Splicing factor 3B	54.5	0.0	22.2
SLC25A	ADP/ATP translocase 2 (solute carrier family 25)	100.0	71.4	100.0
THOC4	THO complex subunit 4	44.4	9.1	0
TKT	Transketolase	36.4	42.9	33.3
TMPO	Thymopoietin	22.2	0	28.6
TOMM22	Translocase of outer membrane 22-kD subunit homologue	22.2	0	28.6
TRAP1	Tumor necrosis factor type 1 receptor-associated protein (heat shock protein 75)	44.4	27.3	42.9
TRIM21	52-kD Ro protein	18.2	28.6	33.3
TRIM25	Tripartite motif-containing protein 25 (Zinc finger protein 147)	18.2	0.0	0.0
TTBK2	Tau-tubulin kinase	18.2	14.3	44.4
TUFM	Tu translation elongation factor, mitochondrial	63.6	14.3	55.6
TXN	Thioredoxin	27.3	57.1	55.6
U2AF1	U2 small nuclear RNA auxiliary factor 1	27.3	0.0	11.1
UBA52	Ubiquitin and ribosomal protein L40 precursor	36.4	57.1	55.6
UBE2D2,3	Ubiquitin-conjugating enzyme E2D 2, E2D 3	18.2	14.3	0.0
UQCRC1	Ubiquinol-cytochrome c reductase core protein I	63.6	0.0	22.2

Table II. **Sepharose bead proteome: other proteins of additional classes** (Continued)

Gene Name	Description	NP	CP	WC
		% Exp.	% Exp.	% Exp.
VAPA/B	VAMP (vesicle-associated membrane protein)-associated protein A, B	44.4	27.3	28.6
VCP	Valosin-containing protein	44.4	18.2	42.9
VDAC2,3	Voltage-dependent anion channel 2, 3	63.6	0.0	22.2
XPO1	Exportin 1	9.1	0.0	22.2
XRCC5, 6	ATP-dependent DNA helicase II	18.2	14.3	0.0
YBX1	Y box binding protein 1	36.4	28.6	33.3
YWHAZ, YWHAB	Tyrosine 3-monooxygenase/tryptophan 5-monooxygenase activation protein	27.3	28.6	44.4

In addition to the common classes of proteins listed in Table I, there are over 100 proteins, shown here, that do not fall into these specific classes yet still bind Sepharose nonspecifically. This list was screened against datasets from 27 independent SILAC immunoprecipitation experiments using either nucleoplasmic (NP; 11 experiments), cytoplasmic (CP; 7 experiments), or whole cell (WC; 9 experiments) extracts to determine the frequency of detection and distribution of these nonspecific binding proteins among these distinct cellular extracts. The frequency is listed as the percentage of experiments in which the protein was detected.

also found by SILAC as a specific interaction partner only in nuclear extract (Fig. 4 D and Table III). Furthermore, the cytoplasm-specific interaction partner PRMT5 was also found here as a specific interaction partner only in cytoplasmic extract (Fig. 4 C and Table III). These results demonstrate the effectiveness of the SILAC approach for identifying specific protein binding partners and show that it can resolve compartment-specific interactions.

Almost all of the other previously reported SMN interaction partners were also found in this analysis (see Table III), although in some cases the SILAC ratios were close to those for nonspecific Sepharose-binding contaminant proteins. The analysis of the SMN complex thus illustrates the importance of including information from additional data to the SILAC ratios, including peptide abundance and bead proteome information, to help distinguish specificity where SILAC ratios are close to background levels. For example, both PRMT5 and Unrip, which have been reported to interact with SMN, show relatively low SILAC ratios compared with the gemins. However, the fact that neither of these proteins was detected binding nonspecifically to either GFP or Sepharose increases the probability that they are specific binders. In contrast, certain proteins with higher SILAC ratios, such as desmin and transketolase, were commonly found in the Sepharose bead proteome, which reduces the probability that they represent specific binding partners for SMN. Peptides were also found for hnRNP Q and RNA helicase A, both reported to interact with SMN (Mourelatos et al., 2001; Pellizzoni et al., 2001b; Rossoll et al., 2002). The peptides were not quantifiable, however, and we therefore did not include them in the list of unambiguously identified known SMN interaction partners. Interestingly, U1 70k protein was found to copurify with GFP-SMN from cytoplasmic extracts, with 15 separate peptides detected with high SILAC ratios. SMN was reported to bind the U1 snRNA and the U1 snRNP-specific A protein, although this interaction with the U1-specific 70k protein was not previously detected (Pellizzoni et al., 2002b).

We have developed a useful strategy for analyzing the SILAC data to help distinguish specific interactions (Fig. 5). Data acquired from SILAC-based quantitative immunoprecipitation experiments are first plotted in a histogram. This helps to visualize the grouping of nonspecific binding proteins, which generally fall within a bell-shaped curve regardless of the abso-

lute value of the SILAC ratios. Although under ideal conditions a ratio of 1 should be obtained for nonspecific binding, this absolute value can vary experimentally in either direction. This is illustrated in Fig. 5 A, where the absolute peak values for the bell-shaped curves for the separate nuclear and cytoplasmic extracts differ slightly. Within each experiment, the SILAC ratios can thus be evaluated with respect to the actual background ratio curve determined and a corresponding threshold set for that experiment (Fig. 5 A, hashed blue and red lines).

To further extend this analysis and improve confidence, the bead proteome data are next applied as a filter to highlight proteins that are known to bind nonspecifically to the affinity matrix and reveal proteins that may bind specifically yet are close to or below the chosen threshold. As illustrated for the cytoplasmic extract, SILAC ratios are first plotted for all proteins previously identified as binding nonspecifically to Sepharose (Fig. 5 B). Proteins that may bind to the GFP tag itself (Fig. 2 C) are also included in this list (Fig. 5 B, green). In the case of hnRNP proteins, which are commonly found in the Sepharose bead proteome, multiple members of the hnRNP family seen in the analysis of SMN-associated proteins are identified as likely contaminants with SILAC ratios at or below the threshold level. However, hnRNP U alone stands out with a higher SILAC ratio in both nuclear and cytoplasmic experiments, consistent with previous evidence reporting hnRNP U as a specific component of the SMN complex (Liu and Dreyfuss, 1996). This demonstrates that not all proteins in the bead proteome are inevitably binding nonspecifically and therefore they should not be excluded on this basis alone from further analysis.

Although the majority of potential contaminants have SILAC ratios either at or near the chosen threshold, some show significantly higher ratios, such as desmin and transketolase. This is either due to a real interaction with GFP-SMN, or to variability inherent in the experiment or in the quantitation. Importantly, by highlighting these proteins as potential contaminants, they may be considered lower priority for future detailed analysis.

Next, filtering out proteins known to bind nonspecifically to Sepharose leaves a list of putative interacting partners that can also be analyzed separately (Fig. 5 C). As shown here, over two-thirds of these proteins have a SILAC ratio

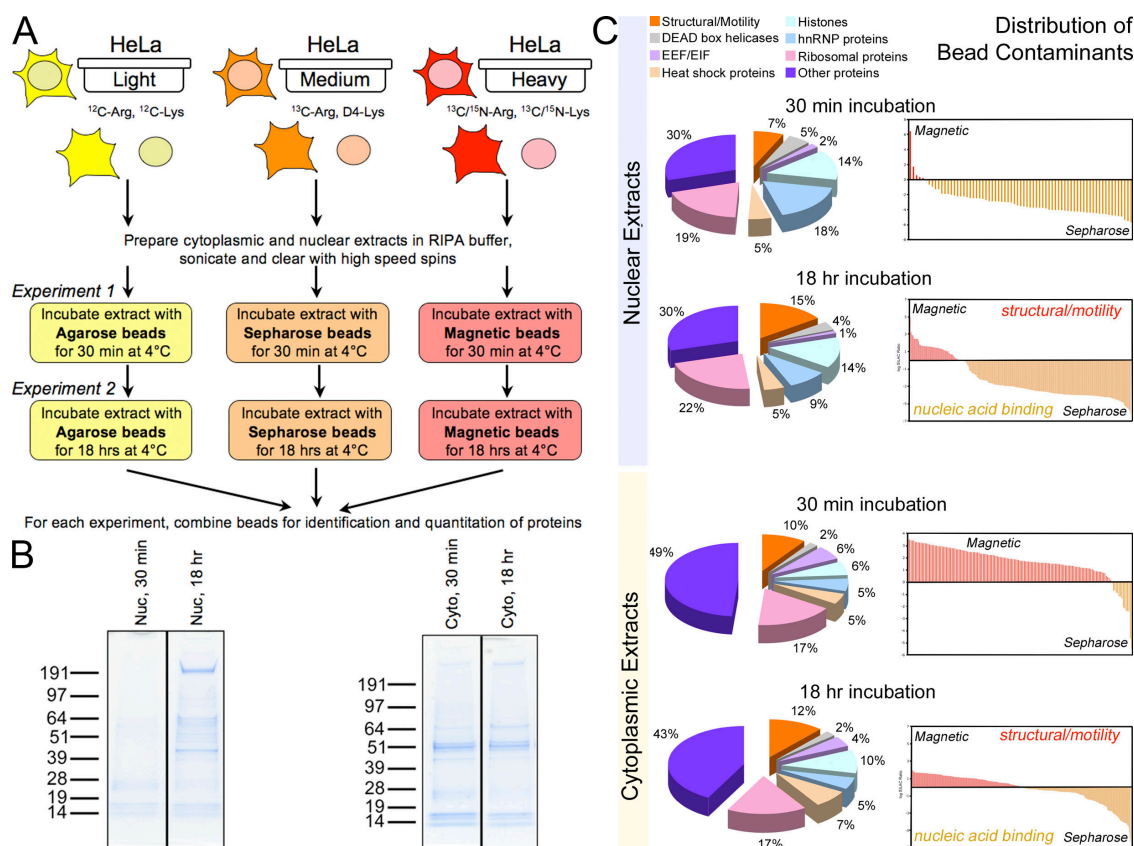


Figure 3. Comparison of bead proteomes. (A) Design of the SILAC immunoprecipitation experiment used to compare the bead proteomes of agarose, Sepharose, and magnetic beads. For all three, the protein G-conjugated versions were used. The experiment was performed in two stages, first with a short incubation time of 30 min and next with a long incubation time of 18 h. In addition, cells were fractionated into cytoplasmic and nuclear extracts to compare the profiles of the proteins that bind nonspecifically to the bead matrices. In the case of nuclear extracts, more proteins bind nonspecifically during a long incubation than a short incubation, as assessed both by Coomassie staining (B) and by mass spectrometric analysis (C). The cytoplasmic protein profile did not vary to the same extent. The distribution of proteins by class was quite similar regardless of the cellular extract used in the experiment or the time of incubation (C). Distinct differences in the distribution of these classes of proteins were observed, however, with magnetic beads binding more cytoskeletal and structural proteins nonspecifically and Sepharose binding more nucleic acid binding factors nonspecifically.

sufficiently high to indicate specific interaction with GFP-SMN, and indeed most are known SMN interaction partners, as detailed in Table III. Of the remaining proteins, several are known SMN interacting partners that, in this experiment, have SILAC ratios close to threshold and thus may have been overlooked in the initial analysis (e.g., Sm proteins, PRMT5, and Unrip). This emphasizes the importance of the enhanced workflow for highlighting specific interaction partners among a sea of contaminants.

Most of the remaining proteins shown in Fig. 5 C have low SILAC ratios and correspond to metabolic enzymes, which at this stage appear as low priority targets for further analysis. However, one of the remaining novel proteins identified here, USP9X, had a higher SILAC ratio (Fig. 5 C) and is known to be a de-ubiquitinating enzyme that was recently shown to regulate AMPK-related kinases (Al-Hakim et al., 2008). We therefore selected this as the highest priority for follow up analysis.

Validation of USP9X by Western blotting

To test whether the identification of USP9X by SILAC analysis can be verified by an independent method, we next performed Western blotting analysis on protein complexes affinity purified

with GFP binder from both cytoplasmic and nuclear extracts (Fig. 6). In this case, cells expressing free GFP were used as a control. An antibody specific to USP9X detected specific pull-down of USP9X by GFP-SMN, especially in the cytoplasmic extracts (Fig. 6 A). This confirms the identification of USP9X in the previous SILAC experiments, and is consistent with the fact that USP9X peptides were only identified by SILAC in the cytoplasmic extract (for an example of a mass spectrum for a USP9X SILAC peptide, see Fig. 6 B). The predominantly cytoplasmic signal of USP9X is also consistent with immunofluorescence analysis. Thus, immunostaining of HeLa cells with anti-USP9X antibody revealed that it is enriched in the cytoplasm, although a weak nucleoplasmic pool is also detected (Fig. 6 C). The localization of endogenous USP9X is the same in the presence (bottom cell) and absence (top cell) of GFP-SMN, and in both cases there is no apparent accumulation in gems. The fact that USP9X had not been identified previously as associating with this well-characterized protein complex suggests that it may either be low abundance, interact transiently with the SMN complex, and/or bind with low affinity.

As a positive control, Western blotting was also performed to confirm the enrichment of SMN and U1A under the same

Table III. Previously reported SMN interaction partners identified in quantitative proteomics screen utilizing GFP binder to immunoprecipitate GFP-SMN complexes

	MW (kD)	Cytoplasmic SILAC Ratio ± SD	# Peptides	Nucleoplasmic SILAC Ratio ± SD	# Peptides	Reference
SMN	31.8	14.0 ± 19.0	15	76.4 ± 96.5	13	(Kolb et al., 2007; Otter et al., 2007)
Gemin 2	31.6	10.8 ± 2.2	10	39.3 ± 26.8	6	(Liu et al., 1997)
Gemin 3	92.2	74.2 ± 87.0	36	33.2 ± 35.6	33	(Charroux et al., 1999; Campbell et al., 2000)
Gemin 4	120	66.5 ± 79.5	54	36.3 ± 46.8	46	(Charroux et al., 2000)
Gemin 5	168.6	5.2 ± 1.3	26	9.9 ± 7.3	19	(Gubitz et al., 2002)
Gemin 6	18.8	50.5 ± 59.8	7	34.3 ± 43.5	6	(Pellizzoni et al., 2002a)
Gemin 7	14.5	36.8 ± 3.2	6	18 ± 12.1	5	(Baccon et al., 2002)
Gemin 8	40.1	80.3 ± 102.3	11	26.1 ± 41.6	9	(Carissimi et al., 2006)
U1A	31.3	3.5 ± 0.3	5	2.5 ± 0.2	4	(Yong et al., 2002)
SmB/B'	30	0.8 ± 0.01	5	1.9 ± 0.1	5	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmD1	13.3	0.9 ± 0.1	6	11.4 ± 2.2	4	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmD2	13.5	0.9 ± 0.1	10	8.9 ± 1.3	7	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmD3	13.9	1.0 ± 0.1	4	2.0 ± 0.2	5	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmE	10.8	1.0 ± 0.02	4	6.5 ± 0.6	4	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmF	9.7	1.0 ± 0.1	3	5.6 ± 0.03	4	(Liu et al., 1997; Pellizzoni et al., 2002b)
SmG	8.5	1.0 ± 0.04	6	6.6 ± 1.4	4	(Liu et al., 1997; Pellizzoni et al., 2002b)
Lsm10	14.1	5.8 ± 2.5	7	7.8 ± 2.7	5	(Pillai et al., 2001, 2003)
Lsm11	39.5	6.7 ± 1.7	11	0.7 ± 0.1	1	(Pillai et al., 2003)
Unrip	38.4	1.4 ± 0.1	11	3.6 ± 0.9	17	(Carissimi et al., 2005)
Coilin	62.6	-	-	21.6 ± 13.0	14	(Hebert et al., 2001)
PRMT5	72.7	1.5 ± 0.3	5	-	-	(Meister and Fischer, 2002)
Fibrillarin	33.8	-	-	2.8 ± 0.5	1	(Liu and Dreyfuss, 1996; Jones et al., 2001; Pellizzoni et al., 2001a)
hnRNP U	90.6	1.7 ± 0.1	8	3.9 ± 0.8	13	(Liu and Dreyfuss, 1996)

This table summarizes the protein interaction datasets collected from both cytoplasmic and nuclear extracts using the GFP binder to pull down GFP-SMN from stable HeLa cell lines. Both the SILAC ratio (with SD) and the number of peptides identified for each protein in a particular experiment are indicated. Ratios >10, which inevitably show higher standard deviations, as discussed in the text, are in bold. References for the initial characterization of each protein as an SMN interaction partner are also listed.

affinity purification conditions in both cytoplasmic and nuclear extracts, and the nuclear extract-specific enrichment of coilin (Fig. 6 D). For comparison, sample mass spectra for SMN, U1A, and coilin peptides identified by SILAC analysis are shown (Fig. 6 E). Although high SILAC ratios reliably distinguish binding specificity, we note that the absolute SILAC ratio cannot currently be used to infer stoichiometry of binding. As shown by the high standard deviation values measured for high SILAC ratios (see Table III, ratios >10 in bold), it is difficult to accurately quantitate ratio values when one of the components used to generate the ratio is present in very low amounts (see representative peptide spectra in Fig. 6 E).

After confirming the positive identification of USP9X, we also tested by Western blotting other proteins that had high SILAC ratios yet were considered more likely to be contaminants based on the SILAC workflow. For example, both desmin and transketolase had high SILAC ratios in the cytoplasmic extract (Fig. 5 B), but did not show specific pull-down as judged by Western blotting (unpublished data). This confirms that they were indeed contaminants, most likely binding nonspecifically to Sepharose beads.

SILAC analysis by direct immunoprecipitation

Finally, we also evaluated the SILAC method using direct immunoprecipitation with an antibody specific for the endogenous

SMN protein. This is important because not all proteins are either functional or correctly expressed after tagging with GFP, and we thus wanted to test whether a similar workflow could be applied for identification of protein partners using antibodies to endogenous proteins. For these experiments we used a monoclonal anti-SMN antibody (BD Biosciences), which was tested and found to specifically immunoprecipitate SMN (see Fig. S3 and Table S2, available at <http://www.jcb.org/cgi/content/full/jcb.200805092/DC1>). A similar overall workflow was applied, with minor modifications (see Fig. 1 B). SILAC analysis of the immunoprecipitated proteins again identified many of the core SMN complex proteins, although the number of peptides and overall quality of the data were notably poorer than that obtained using the GFP binder and GFP-tagged SMN (Fig. S3 and Table S2). One reason for this is likely the less efficient depletion of endogenous SMN by the anti-SMN mAb as compared with the near-quantitative depletion of GFP-SMN using the GFP binder. It appears this is not simply a question of overall expression levels, however, as GFP-SMN is expressed in the stable cell line at a lower level than endogenous SMN (Sleeman et al., 2003). To test this idea, we compared the data resulting from pull-down of GFP-SMN using the GFP binder with a pull-down using the commercial anti-GFP mAb previously shown to be less efficient in depletion of GFP (see Fig. 2). The quality of the resulting data, including the number

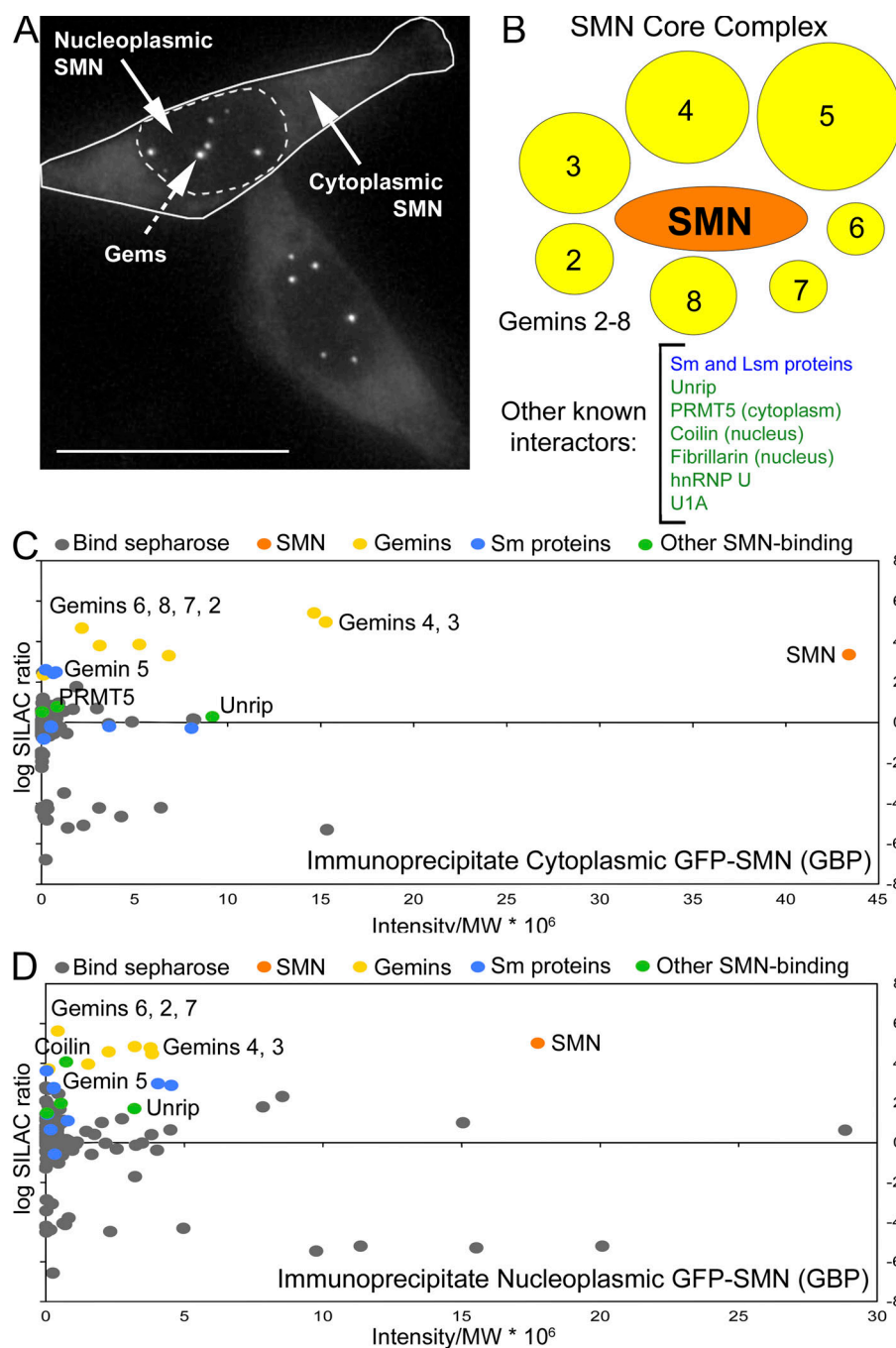


Figure 4. Identification of proteins that interact with SMN and the SMN complex. The GFP binder was used to immunopurify GFP-SMN from a stable HeLa cell line as compared with the nonexpressing parental cell line. Like endogenous SMN, GFP-SMN is found in both cytoplasmic and nucleoplasmic pools and accumulates in gems within nuclei (A). Bar, 15 μ M. Detailed biochemical and proteomic studies have revealed that the core SMN complex is composed of SMN itself and Gemin 2–8 (B). The stoichiometry is not known and, although not depicted here, the complex can oligomerize. Also listed are several other proteins that have been shown to interact with the SMN complex by similar experimental approaches. In the study presented here, separate experiments were performed for cytoplasmic and nuclear extracts to independently assess interacting partners and compare these two pools. The log SILAC (i.e., heavy/light arginine and/or lysine) ratio calculated for each protein identified in the cytoplasmic GFP-SMN immunoprecipitation experiment is plotted versus total peptide intensity in C. The nucleoplasmic GFP-SMN immunoprecipitation data are plotted in a similar fashion (D).

of peptides identified and quantified, was clearly better using the GFP binder as compared with the commercial anti-GFP mAb (Fig. S3; Table S2).

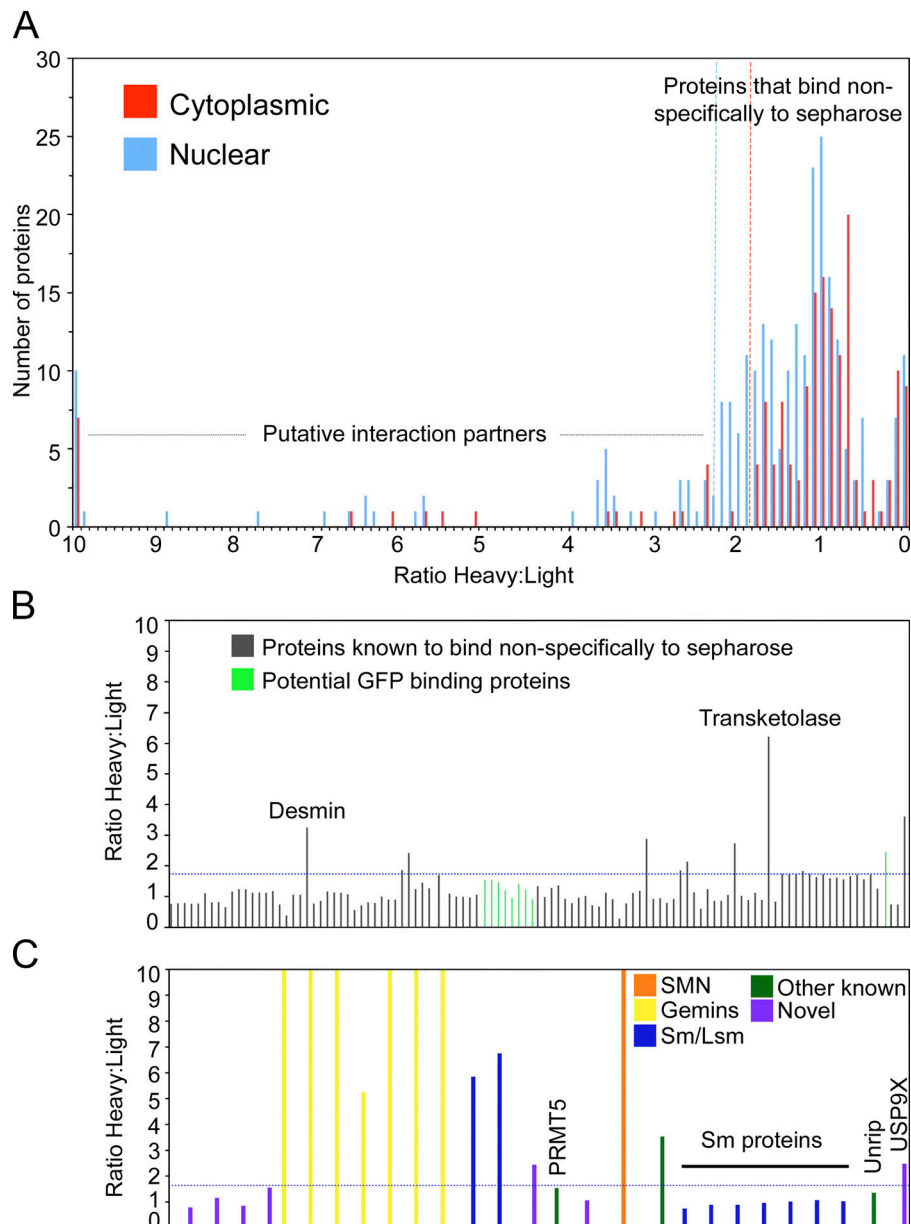
In summary, these data show that the SILAC approach can be successfully applied for the analysis of endogenous proteins directly immunoprecipitated with antibodies. However, the overall quality of the resulting data will inevitably be affected by the specificity and efficiency of the available antibodies.

Discussion

This study describes a method based on quantitative SILAC mass spectrometry (Ong et al., 2002) that has been optimized to

facilitate the reliable detection of bona fide protein interaction partners in cell extracts by immuno- and/or affinity purification. This approach has been made possible thanks to the recent major advances in the sensitivity and mass accuracy of mass spectrometry-based proteomics (Domon and Aebersold, 2006; Cox and Mann, 2007). These technological improvements facilitate detection of lower abundance proteins and allow for a genuine high-throughput approach. Increased sensitivity of detection alone does not reliably identify specific interaction partners, however, as there is a concomitant detection also of the many nonspecifically bound proteins that routinely copurify in pull-down experiments. To minimize contaminants, many previous studies have used high stringency purification methods. This is

Figure 5. Systematic analysis of SILAC data-sets. Quantitative mass spectrometric data generated by the cytoplasmic and nuclear GFP-SMN immunoprecipitation experiments were subjected to a standard analysis workflow. First, the frequency of specific SILAC (heavy/light amino acid) ratios were plotted for the entire datasets to determine the distribution of these ratios among the proteins identified (A). Environmental contaminants such as keratins have very low ratios and cluster near 0. In the cytoplasmic experiment, proteins that bind non-specifically to the bead matrix cluster in a bell curve distribution around 1, as expected for proteins that bind equally in the light and heavy form. The threshold for detection of bona fide interaction partners was set at a conservative level above that (hashed red line). Note that in the nuclear experiment the SILAC ratios for the bead contaminants were shifted to the left, clustering in a bell curve distribution around the higher value of 1.5. In this case the threshold (hashed blue line) must also be shifted. SMN itself, all of the core SMN complex members, and several known interacting partners fell above this threshold and were identified in this first analysis step. However, less abundant or lower affinity binding partners may be found at or below these conservative threshold values. Analysis of the datasets is thus further extended by applying the Sepharose bead proteome as a filter and grouping the SILAC ratios of those proteins that have been identified as binding nonspecifically to this bead matrix, as shown here for the cytoplasmic dataset (B). Most proteins known to bind Sepharose (gray) and potential GFP-binding proteins (green) have the expected ratios near or below threshold, but a few are significantly above threshold and must be considered as potentially real interacting proteins, albeit with a lower priority for further analysis. SILAC ratios calculated for the remaining proteins in the dataset, i.e., those not known to bind nonspecifically to either the GFP tag or the bead matrix, are next plotted separately (C). Over two-thirds of the proteins have SILAC ratios significantly higher than threshold. These include both known and novel interacting partners for SMN. Some of the known SMN complex interacting partners, such as PRMT5 and Unrip, have ratios closer to threshold, and thus would be overlooked in a threshold-based analysis. As expected for such a well-characterized complex, very few novel proteins were detected. One of these, USP9X, was selected for further analysis.



also not ideal because stringent purification procedures often result in the loss of specific binding partners, for example those interacting in sub-stoichiometric amounts or binding with lower affinity. The strategy described here takes advantage of the sensitivity of modern mass spectrometry-based proteomics to identify en masse components of protein complexes purified under lower stringency conditions, which preserves more specific interactions.

A key feature of the method involves combining SILAC ratios with bead proteomes and other data filtering to distinguish likely specific interacting proteins from the much larger pool of nonspecific binding proteins (see Fig. 5). This is particularly valuable in assessing whether proteins with SILAC ratios

close to threshold values represent specific interaction partners. This strategy can be applied directly to analyze endogenous protein complexes isolated by immunoprecipitation. In addition, we show that it can provide a powerful dual strategy when applied to the analysis of proteins interacting with GFP-tagged fusion proteins in a “what you see is what you get” approach. Importantly, this allows the integration of biochemical in vitro information derived from analysis of pull-down experiments, with in vivo data describing the localization, dynamics, and protein interactions derived from fluorescence microscopy. In contrast, the use of separate tags for affinity purification studies and microscopy analysis does not allow a direct comparison of the data obtained. GFP has been used previously as an affinity tag

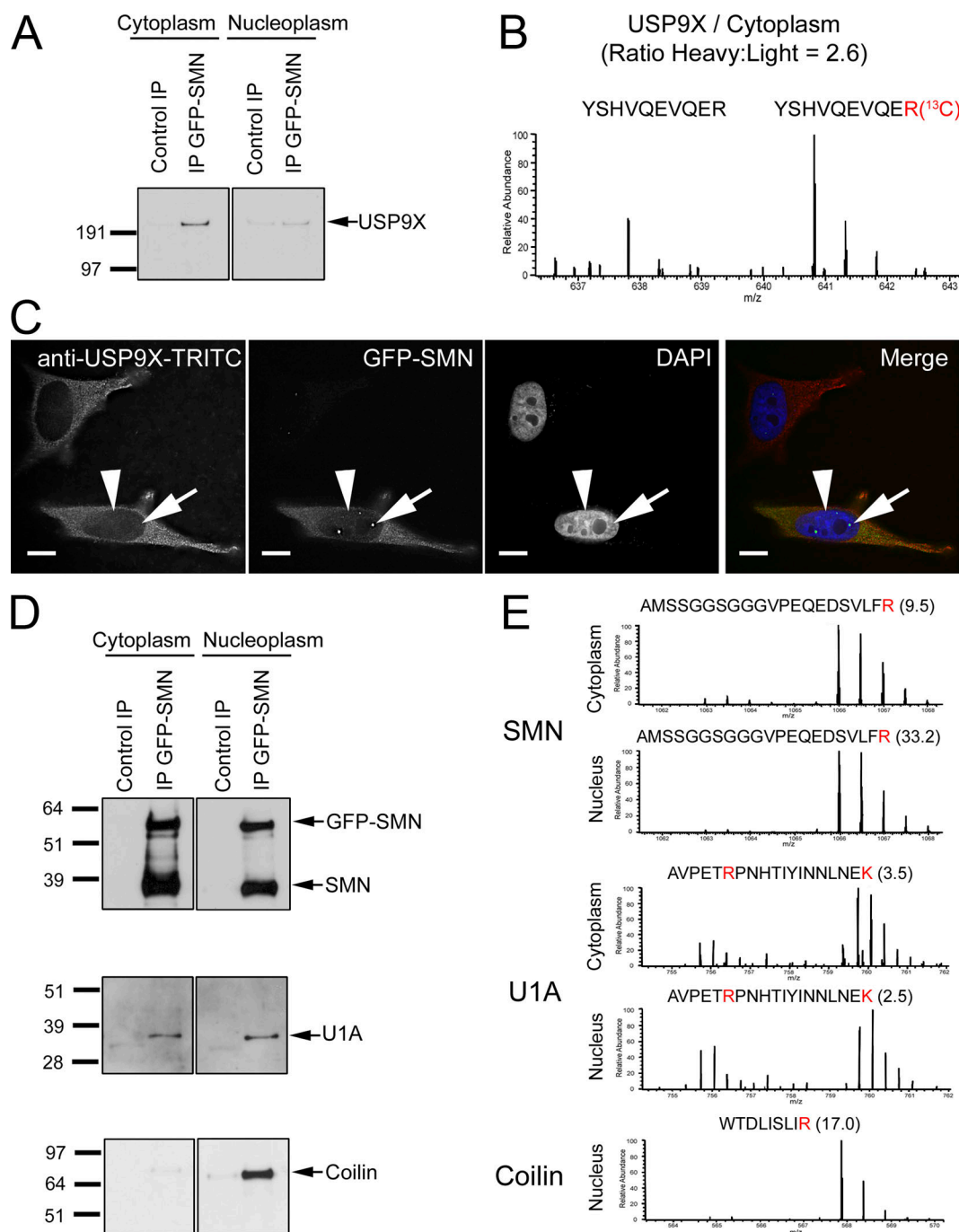


Figure 6. Validation of mass spectrometric results. Cytoplasmic-specific copurification of the novel protein USP9X with GFP-SMN was confirmed by Western blotting (A). Two peptides, each with a SILAC ratio >1, were found for USP9X in the SILAC analysis of a GFP-SMN pull-down from cytoplasmic extracts. The mass spectra of one of them is shown here for comparison (B). The quantifiable arginine is highlighted in red. This cytoplasmic enrichment of USP9X is consistent with immunostaining results using a monoclonal anti-USP9X antibody (C). Although predominantly cytoplasmic, there is a pool of USP9X in the nucleus (arrowhead), although it does not accumulate in gems (arrow). There is no difference in localization of USP9X in parental HeLa cells (top cell) versus HeLa cells stably expressing GFP-SMN (bottom cell). Bar, 5 μ M. As a control, Western blotting was also used to confirm the enrichment of both endogenous SMN and GFP-SMN, and of the U1 snRNP protein U1A, from both cytoplasmic and nuclear extracts using the GFP binder, and the nuclear-specific enrichment of p80 coilin (D). For comparison, representative peptide spectra for these proteins from the SILAC analysis are shown (E). Quantifiable amino acids are highlighted in red, with the SILAC ratio in parentheses.

for proteomics studies (Cristea et al., 2005; Trinkle-Mulcahy et al., 2006). The results in this study underline the suitability of GFP as a dual strategy tag. First, both in vivo photobleaching experiments and SILAC mass spectrometry data show that GFP exhibits minimal nonspecific binding to mammalian cell pro-

teins. Second, the recent advent of the GFP binder affinity probe allows near-quantitative depletion of GFP fusion proteins from cell extracts, thereby improving signal-to-noise ratios and maximizing the range of protein complexes that can be recovered. Based on the successful analysis of over 20 separate GFP fusion

proteins in whole cell, cytoplasmic, and nuclear extracts, our results indicate that a similar strategy can be readily applied for the analysis of interaction partners binding to most, if not all, GFP-tagged proteins.

In the SILAC-based strategy for analyzing protein interaction partners (see Fig. 1), the ratio of heavy to light isotopes measured for each peptide detected provides an unbiased and often clear-cut index for distinguishing specific from nonspecific binding proteins (for examples of peptide spectra, see Fig. 6). In some cases, however, particularly for lower abundance proteins, the $^{13}\text{C}/^{12}\text{C}$ (SILAC) ratio alone is not sufficient to unambiguously distinguish specificity. The order of steps in the workflow and the detailed experimental protocol can be sources of variability. For example, accurately controlling the amounts of material mixed together before or after immunoprecipitation can affect the ratio. In addition, the ratio can also be affected by dissociation of proteins from the complex during isolation. Depending on the complex under study, it could also happen that exchange occurs between the isotope-labeled proteins on the affinity matrix and proteins in the control extract (Wang and Huang, 2008). For these reasons, our results show it is important to minimize the binding time whenever possible, which will also help to reduce the level of nonspecific protein binding. This latter point is illustrated by the larger cohort of nonspecific binding proteins recovered after extended (18 h) incubation of the extracts with all three affinity matrices (see Fig. S2 and Table S1). Finally, it is also important to optimize the efficiency of protein pull-down. This is best illustrated by the comparison of using a commercial anti-GFP mAb as compared with GFP binder to affinity purify GFP-SMN (see Fig. S3 and Table S2).

As illustrated here by the analysis of the well-characterized SMN complex, a useful additional criterion to add to the SILAC ratio is to filter all identified proteins against a database of proteins found to bind nonspecifically to affinity matrices under a range of conditions. This was shown to help distinguish known SMN interaction partners from likely contaminants (see Table III and Fig. 5). In the case of Sepharose, the bead proteome was derived from 27 different SILAC-based pull-down experiments. This includes separate analysis for pull-downs performed in whole cell, nuclear, and cytoplasmic extracts for both HeLa and U2OS cell lines. Identical results were obtained for both cell lines and the data have therefore been combined in the Sepharose bead proteome presented (Tables I and II). Interestingly, similar sets of protein contaminants were identified in the separate cytoplasmic and nuclear extracts, including ribosomal, heat shock, hnRNP, and intermediate filament proteins. We extended the analysis of the bead proteome to include direct comparisons of Sepharose, agarose, and magnetic beads, which to the best of our knowledge currently represent the three most commonly used affinity matrices. Unexpectedly, differences were observed in the spectrum of contaminating proteins that predominate for each of these matrices, and this varied between the separate nuclear and cytoplasmic extracts. Thus, we did not observe a single bead matrix that gave universally lower levels of contaminants under all circumstances. For cytoplasmic extracts, the lowest background levels were obtained using either Sepharose or agarose. Magnetic beads, in contrast, showed more nonspecific binding

for cytoskeletal and structural proteins that are abundant in cytoplasmic extracts. Conversely, magnetic beads showed lower nonspecific binding to nucleic acid-associated proteins and thus gave lower backgrounds than either Sepharose or agarose when used with nuclear extracts. These data provide objective grounds for concluding that no single type of affinity matrix is best for all purposes, and highlights the importance of choosing the most suitable combination of reagents based on the specific details of the experiment to be performed.

An important question raised by this identification of many proteins that clearly bind nonspecifically to commonly used affinity matrices in protein–protein interaction experiments is the accuracy of the published literature. In many cases, published studies have listed as potential interaction partners proteins shown here to bind nonspecifically to affinity matrices. The bead proteome filters thus provide a useful and objective resource that can be consulted by cell biologists to help avoid expending time and effort on the analysis of proteins that may prove to be simple contaminants. In the future, accumulating information from many laboratories on the range of nonspecific protein interactions observed using different cell types, extracts, tags, and affinity matrices will provide an invaluable resource and we propose this should be established as a freely accessible online database.

In summary, the present data show that a strategy combining SILAC analysis with bead proteome filtering and enhanced data analysis procedures can reliably be used to characterize specific protein interaction partners while using isolation procedures that preserve the binding of lower abundance and lower affinity proteins. We show that this can also resolve interaction events confined to either nuclear or cytoplasmic compartments. Inevitable differences in the biochemical properties of different proteins mean that no unique isolation protocol may be ideal in every case. Nonetheless, we could show that a similar isolation protocol could be successfully applied to analyze over 20 different GFP fusion proteins in multiple different cell extracts and from two separate mammalian cell lines. Even when precise isolation conditions must be varied, our data indicate general principles that apply, including the importance of maintaining short incubation times during affinity purification and the need to optimize the overall efficiency of affinity depletion. We show the strategy can be used for the analysis of tagged or endogenous complexes and thus conclude it provides a general approach that can be widely applied for the analysis of protein binding partners in different fields of cell biology.

Materials and methods

Tissue culture

HeLa^{EGFP} and HeLa^{EGFP-SMN} stable cell lines were obtained and characterized as described previously (Sleeman et al., 2003). Cells were grown in custom-made DMEM (minus arginine and lysine; Invitrogen) supplemented with 10% dialyzed fetal calf serum (Invitrogen) and penicillin/streptomycin (Invitrogen). The selection marker G418 was added to SILAC media used with stable cell lines expressing GFP-tagged proteins. For double encoding experiments, L-arginine (84 µg/ml; Sigma-Aldrich) and L-lysine (146 µg/ml lysine; Sigma-Aldrich) were added to the “light” media, while L-arginine ^{13}C and L-lysine ^{13}C (Cambridge Isotope Laboratory) were added to the “heavy” media at the same concentrations. For triple encoding experiments, L-arginine and L-lysine were added to the “light”, L-arginine ^{13}C and L-lysine 4,4,5,5-D4 (Cambridge Isotope Laboratory) to

the "medium", and L-arginine $^{13}\text{C}/^{15}\text{N}$ and L-lysine $^{13}\text{C}/^{15}\text{N}$ (Cambridge Isotope Laboratory) to the "heavy" media. The amino acid concentrations are based on the formula for normal DMEM (Invitrogen). Once prepared, the SILAC media was mixed well, filtered through a 0.22- μm filter (Millipore) using a suction pump, and stored at 4°C. HeLa and U2OS cell lines were passaged in SILAC media for at least 5–6 cell doublings before harvesting to ensure complete incorporation of isotopic amino acids (Ong and Mann, 2007; Harsha et al., 2008). PBS-based nonenzymatic cell dissociation buffer (Invitrogen) was used to passage cells, as trypsin-EDTA solutions may contain amino acids.

Preparation of cellular extracts

Whole cell extracts were prepared by solubilizing trypsinized and pelleted cells in ice-cold RIPA buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1% NP-40, 0.5% deoxycholate, and protease inhibitors), sonicating briefly on ice (5 × 10 s at full power), and clearing extracts by centrifuging at 2,800 g (3,500 rpm, GH3.8 rotor; Beckman Coulter GS-6) for 10 min at 4°C. For preparation of cytoplasmic and nuclear fractions, 10 × 14-cm dishes of cells were trypsinized and pelleted, resuspended in 5 ml of ice-cold swelling buffer (10 mM Hepes, pH 7.9, 1.5 mM MgCl_2 , 10 mM KCl, 0.5 mM DTT, and protease inhibitors) for 5 min, and cells were broken open to release nuclei using a pre-chilled Dounce homogenizer (20 strokes with a tight pestle). Dounced cells were centrifuged at 228 g (1,000 rpm, GH-3.8 rotor; Beckman Coulter GS-6) for 5 min at 4°C to pellet nuclei and other fragments. The supernatant was retained as the cytoplasmic fraction. Before use, 1 ml of 5× RIPA buffer was added and clearing performed as described above. The nuclear pellet was resuspended in 3 ml of 0.25 M sucrose/10 mM MgCl_2 and layered over a 3-ml cushion of 0.88 M sucrose/0.5 mM MgCl_2 and centrifuged at 2,800 g (3,500 rpm, GH-3.8 rotor; Beckman Coulter GS-6) for 10 min at 4°C. The resulting cleaner nuclear pellet was resuspended in 5 ml of RIPA buffer, sonicated and cleared as described above. Total protein concentrations were measured using a Bradford assay.

Immunoaffinity purification of GFP-tagged and endogenous proteins

Monoclonal anti-GFP antibodies (Roche) were covalently coupled to protein G-Sepharose beads (GE Healthcare) at 2 mg/ml. The beads were incubated with antibody for 1 h at 4°C and then washed twice with 10 volumes of 0.1 M sodium borate, pH 9. Next, the beads were incubated with 10 volumes of borate buffer containing 20 mM dimethylpiperimidate (DMP; Sigma-Aldrich) for 30 min at room temperature. The beads were pelleted and resuspended with 10 volumes of freshly prepared 20 mM DMP in borate buffer for an additional 30-min incubation. The beads were washed twice with 10 volumes of ice-cold 50 mM glycine (pH 2.5) to remove unbound antibody and then washed several times with PBS or RIPA buffer for use and/or storage at 4°C. Monoclonal anti-SMN antibodies (BD Biosciences) were covalently coupled to protein G-Sepharose at 1 mg/ml using a similar protocol. GFP binder (ChromoTek) was prepared and covalently coupled to NHS-activated Sepharose 4 Fast Flow beads (GE Healthcare) at 1 mg/ml as described previously (Rothbauer et al., 2008).

For the GFP immunoaffinity experiments, extracts from each cell line were precleared by incubation on Sepharose beads alone for 30 min at 4°C and then mixed in a 1:1 ratio based on total protein concentration. GFP alone or GFP-SMN were affinity purified by incubation with either anti-GFP mAbs or GFP binder conjugated to Sepharose beads. Incubation times varied according to the antibody and the experiment, and we recommend a maximum 1-h incubation, if possible. The affinity matrix was washed four times with RIPA buffer. To ensure efficient elution of bound proteins, a bead-equivalent volume of 1% SDS was added, the matrix boiled for 10 min and then a 4× volume of dH_2O added. The matrix was vortexed and the solution removed and reduced to the original bead-equivalent volume (and 1% SDS concentration) using a speedvac. Proteins were reduced and alkylated in this solution, first by the addition of 10 mM DTT (boil for 2 min), and then the addition of 50 mM iodoacetamide (incubate at room temperature in the dark for 30 min). A small aliquot of Laemmli sample buffer was added and proteins were separated by running halfway down NuPAGE 12% Bis-Tris gels. Gels were Coomassie stained and destained overnight before excision of slices. Peptides resulting from in-gel digestion with trypsin (Promega) were extracted from the gel slices for automated LC-MS/MS analysis. For validation of SILAC results, GFP and GFP-SMN were affinity purified separately using the GFP binder and subjected to 1D SDS/PAGE and Western blotting. Primary antibodies used for Western blotting (and immunofluorescence, where indicated) included anti-USP9X (AbCam mAb, 1:500 WB, 1:50 IF), anti-coilin (204/10 rabbit polyclonal, 1:1,000 WB), anti-SMN (BD Biosciences mAb, 1:1,000 WB),

anti-U1A (856 rabbit polyclonal, 1:2,000 WB), anti-desmin (Abcam mAb; 1:500 WB), and anti-transketolase (goat polyclonal, 1:500; Santa Cruz Biotechnology, Inc.). HRP-conjugated secondary antibodies (Thermo Fisher Scientific) were detected using the ECL-Plus reagent (GE Healthcare).

For the endogenous SMN immunoaffinity experiment and the bead proteome experiment comparing Protein G-Agarose (GE Healthcare), Protein G-Sepharose (GE Healthcare), and the magnetic Protein G-Dynabeads (Invitrogen), equivalent total protein amounts of extracts were incubated separately on the appropriate matrices and combined carefully after one wash step in RIPA buffer. After a further three wash steps in RIPA buffer, bound proteins were eluted and subjected to 1D SDS/PAGE followed by band excision and peptide digestion as described above.

Mass spectrometry and data analysis

An aliquot of the tryptic digest (prepared in 5% acetonitrile/0.1% trifluoroacetic acid in water) was analyzed by LC-MS on an LTQ-Orbitrap mass spectrometer system (ThermoElectron) coupled to a Dionex 3000 nano-LC system (Camberley). The peptide mixture was loaded onto an LC-Packings PepMap C18 column trap column (0.3 × 5 mm) equilibrated in 0.1% TFA in water at 20 $\mu\text{l}/\text{min}$, washed for 3 min at the same flow rate, and then the trap column was switched in-line with an LC-Packings PepMap C18 column (0.075 × 150 mm) equilibrated in 0.1% formic acid/water. The peptides were separated with a 55-min discontinuous gradient of acetonitrile/0.1% formic acid (2–40% acetonitrile for 40 min) at a flow rate of 300 nl/min and the HPLC interfaced to the mass spectrometer with an FS360-20-10 picotip (New Objective) fitted to a nanospray 1 interface (ThermoElectron) with a voltage of 1.1 kV applied to the liquid junction.

The Orbitrap was set to analyze the survey scans at 60,000 resolution and the top five ions in each duty cycle selected for MSMS in the LTQ linear ion trap. The raw files were processed to generate a Mascot generic file using the program Raw2msm (Olsen et al., 2005) and searched against the UniProt human database using the Mascot search engine v.2.2 (Matrix Science) run on an in-house server using the following criteria; peptide tolerance = 10 ppm, trypsin as the enzyme and carboxyamidomethylation of cysteine as a fixed modification. Variable modifications were oxidation of methionine, medium SILAC labels were: label $^{13}\text{C}(6)$ (R), label $^2\text{H}(4)$ (K), and heavy SILAC labels were: label $^{13}\text{C}(6)$ $^{15}\text{N}(4)$ (R), label $^{13}\text{C}(4)$ $^{15}\text{N}(2)$ (K).

Quantitation was performed using the program MS-Quant (<http://msquant.sourceforge.net>), with peptide ratios calculated for each arginine- and/or lysine-containing peptide as the peak area of labeled arginine/lysine divided by the peak area of nonlabeled arginine/lysine for each single-scan mass spectrum. Peptide ratios for all arginine- and lysine-containing peptides sequenced for each protein were averaged. Individual spectra were inspected using QualBrowser software (XCalibur; ThermoElectron). ProteinCenter (Proxeon Bioinformatics) proteomics data mining and management software was used to eliminate redundancy and compare datasets, and to convert protein IDs to gene symbols and perform initial Gene Ontology characterization.

Fluorescence microscopy and photobleaching experiments

Fluorescence imaging was performed on a DeltaVision Spectris widefield deconvolution microscope (Applied Precision) fitted with an environmental chamber (Solent Scientific) to maintain temperature at 37°C, a CoolMax charge-coupled device camera (Roper Scientific) and a quantifiable laser module (QLM; Applied Precision) with a 488-nm laser. For fixed cell imaging, a mix of parental HeLa cells and HeLa cells stably expressing GFP-SMN were paraformaldehyde fixed on glass coverslips, permeabilized with Triton X-100, stained with both anti-USP9X (detected by TRITC-anti-mouse secondary antibodies) and the DNA stain DAPI, and mounted in FluorSave mounting media (Calbiochem). Cells were imaged using a 60× NA 1.4 Plan-Apochromat objective (Olympus) and the appropriate filter sets (Chroma Technology Corp.), with 20 optical sections of 0.5 μm each acquired. SoftWorX software (Applied Precision) was used for both acquisition and deconvolution. For the FRAP experiments, HeLa cells stably expressing free GFP were cultured in glass-bottomed dishes (WILLCO, Intracel) and mounted on the same system. A single section was imaged before photobleaching, a region of interest was then bleached to ~50% of its original intensity using the 488-nm laser, and a rapid series of images was acquired after the photobleach period. Recovery curves were plotted and the mobile fraction and half time of recovery were determined using SoftWorX.

Online supplemental material

Table S1 contains a comprehensive list of all proteins identified in the comparative bead proteome SILAC experiment, including separate datasets for

cytosolic and nuclear extracts and for 30-min and 18-h incubations. Preferential enrichment on either Sepharose or magnetic beads is indicated and commonly found keratins are listed separately. Table S2 compares the quality of data obtained for known SMN complex members using either GFP binder or mAb anti-GFP to affinity purify GFP-SMN and mAb anti-SMN to affinity purify endogenous SMN. Fig. S1 demonstrates the rapid recovery of free GFP in both the cytoplasm and nucleoplasm after photobleaching in live cells. Fig. S2 compares the distribution of nonspecific protein binding between Sepharose and agarose and between magnetic beads and agarose. Fig. S3 graphically compares data obtained using either GFP binder or mAb anti-GFP to affinity purify GFP-SMN and mAb anti-SMN to affinity purify endogenous SMN. Coomassie gels used to separate proteins before mass spectrometric analysis are shown, and SILAC ratio vs. total peptide abundance plotted for known SMN complex members. Online supplemental material is available at <http://www.jcb.org/cgi/content/full/jcb.200805092/DC1>.

We would like to thank Drs. Douglas Lamont and Kenneth Beattie of the Fingerprints Proteomics Facility at the University of Dundee for technical assistance.

Work in the Lamont laboratory was funded by a Wellcome Trust Program Grant (073980/Z/03/Z), and by an interdisciplinary RASOR (Radical Solutions for Researching the Proteome) initiative, which is supported by the Biotechnology and Biological Sciences Research Council, Engineering and Physical Sciences Research Council, Scottish Higher Education Funding Council, and Medical Research Council (MRC). A. Lamont is a Wellcome Trust Principal Research Fellow. S. Boulon is funded by a Human Frontier Science Program fellowship. F.-M. Boisvert is funded by a Caledonian Research Foundation fellowship. N.A. Morrice is supported by the MRC, and F. Vandermoere and R. Urcia are funded by the RASOR collaboration. U. Rothbauer and H. Leonhardt are members of the Munich Center for Integrated Protein Science (CiPSM) and shareholders of ChromoTek (Munich, Germany).

Submitted: 16 May 2008

Accepted: 22 September 2008

References

- Al-Hakim, A.K., A. Zagorska, L. Chapman, M. Deak, M. Pegg, and D.R. Alessi. 2008. Control of AMPK-related kinases by USP9X and atypical Lys(29)/Lys(33)-linked polyubiquitin chains. *Biochem. J.* 411:249–260.
- Baccon, J., L. Pellizzoni, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002. Identification and characterization of Gemin7, a novel component of the survival of motor neuron complex. *J. Biol. Chem.* 277:31957–31962.
- Blagoev, B., I. Kratchmarova, S.E. Ong, M. Nielsen, L.J. Foster, and M. Mann. 2003. A proteomics strategy to elucidate functional protein-protein interactions applied to EGF signaling. *Nat. Biotechnol.* 21:315–318.
- Campbell, L., K.M. Hunter, P. Mohaghegh, J.M. Tinsley, M.A. Brasch, and K.E. Davies. 2000. Direct interaction of Smn with dp103, a putative RNA helicase: a role for Smn in transcription regulation? *Hum. Mol. Genet.* 9:1093–1100.
- Carissimi, C., J. Baccon, M. Straccia, P. Chiarella, A. Maiolica, A. Sawyer, J. Rappsilber, and L. Pellizzoni. 2005. Unrip is a component of SMN complexes active in snRNP assembly. *FEBS Lett.* 579:2348–2354.
- Carissimi, C., L. Saieva, J. Baccon, P. Chiarella, A. Maiolica, A. Sawyer, J. Rappsilber, and L. Pellizzoni. 2006. Gemin8 is a novel component of the survival motor neuron complex and functions in small nuclear ribonucleoprotein assembly. *J. Biol. Chem.* 281:8126–8134.
- Charroux, B., L. Pellizzoni, R.A. Perkinson, A. Shevchenko, M. Mann, and G. Dreyfuss. 1999. Gemin3: A novel DEAD box protein that interacts with SMN, the spinal muscular atrophy gene product, and is a component of gems. *J. Cell Biol.* 147:1181–1194.
- Charroux, B., L. Pellizzoni, R.A. Perkinson, J. Yong, A. Shevchenko, M. Mann, and G. Dreyfuss. 2000. Gemin4. A novel component of the SMN complex that is found in both gems and nucleoli. *J. Cell Biol.* 148:1177–1186.
- Cox, J., and M. Mann. 2007. Is proteomics the new genomics? *Cell.* 130:395–398.
- Cristea, I.M., R. Williams, B.T. Chait, and M.P. Rout. 2005. Fluorescent proteins as proteomic probes. *Mol. Cell. Proteomics.* 4:1933–1941.
- Domon, B., and R. Aebersold. 2006. Mass spectrometry and protein analysis. *Science.* 312:212–217.
- Gubitz, A.K., Z. Mourelatos, L. Abel, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002. Gemin5, a novel WD repeat protein component of the SMN complex that binds Sm proteins. *J. Biol. Chem.* 277:5631–5636.
- Harsha, H.C., H. Molina, and A. Pandey. 2008. Quantitative proteomics using stable isotope labeling with amino acids in cell culture. *Nat. Protoc.* 3:505–516.
- Hebert, M.D., P.W. Szymczyk, K.B. Shpargel, and A.G. Matera. 2001. Coilin forms the bridge between Cajal bodies and SMN, the spinal muscular atrophy protein. *Genes Dev.* 15:2720–2729.
- Jones, K.W., K. Gorzynski, C.M. Hales, U. Fischer, F. Badbanchi, R.M. Terns, and M.P. Terns. 2001. Direct interaction of the spinal muscular atrophy disease protein SMN with the small nucleolar RNA-associated protein fibrillarin. *J. Biol. Chem.* 276:38645–38651.
- Kolb, S.J., D.J. Battle, and G. Dreyfuss. 2007. Molecular functions of the SMN complex. *J. Child Neurol.* 22:990–994.
- Liu, Q., and G. Dreyfuss. 1996. A novel nuclear structure containing the survival of motor neurons protein. *EMBO J.* 15:3555–3565.
- Liu, Q., U. Fischer, F. Wang, and G. Dreyfuss. 1997. The spinal muscular atrophy disease gene product, SMN, and its associated protein SIP1 are in a complex with spliceosomal snRNP proteins. *Cell.* 90:1013–1021.
- Meister, G., and U. Fischer. 2002. Assisted RNP assembly: SMN and PRMT5 complexes cooperate in the formation of spliceosomal UsnRNPs. *EMBO J.* 21:5853–5863.
- Moorhead, G.B., L. Trinkle-Mulcahy, and A. Ulke-Lemee. 2007. Emerging roles of motor protein phosphatases. *Nat. Rev. Mol. Cell Biol.* 8:234–244.
- Mourelatos, Z., L. Abel, J. Yong, N. Kataoka, and G. Dreyfuss. 2001. SMN interacts with a novel family of hnRNP and spliceosomal proteins. *EMBO J.* 20:5443–5452.
- Mousson, F., A. Kolkman, W.W. Pijnappel, H.T. Timmers, and A.J. Heck. 2008. Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol. Cell. Proteomics.* 7:845–852.
- Olsen, J.V., L.M. de Godoy, G. Li, B. Macek, P. Mortensen, R. Pesch, A. Makarov, O. Lange, S. Horning, and M. Mann. 2005. Parts per million mass accuracy on an Orbitrap mass spectrometer via lock mass injection into a C-trap. *Mol. Cell. Proteomics.* 4:2010–2021.
- Ong, S.E., and M. Mann. 2007. Stable isotope labeling by amino acids in cell culture for quantitative proteomics. *Methods Mol. Biol.* 359:37–52.
- Ong, S.E., B. Blagoev, I. Kratchmarova, D.B. Kristensen, H. Steen, A. Pandey, and M. Mann. 2002. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics.* 1:376–386.
- Otter, S., M. Grimmmler, N. Neuenkirchen, A. Chari, A. Sickmann, and U. Fischer. 2007. A comprehensive interaction map of the human survival of motor neuron (SMN) complex. *J. Biol. Chem.* 282:5825–5833.
- Pellizzoni, L., J. Baccon, B. Charroux, and G. Dreyfuss. 2001a. The survival of motor neurons (SMN) protein interacts with the snRNP proteins fibrillarin and GAR1. *Curr. Biol.* 11:1079–1088.
- Pellizzoni, L., B. Charroux, J. Rappsilber, M. Mann, and G. Dreyfuss. 2001b. A functional interaction between the survival motor neuron complex and RNA polymerase II. *J. Cell Biol.* 152:75–85.
- Pellizzoni, L., J. Baccon, J. Rappsilber, M. Mann, and G. Dreyfuss. 2002a. Purification of native survival of motor neurons complexes and identification of Gemin6 as a novel component. *J. Biol. Chem.* 277:7540–7545.
- Pellizzoni, L., J. Yong, and G. Dreyfuss. 2002b. Essential role for the SMN complex in the specificity of snRNP assembly. *Science.* 298:1775–1779.
- Pillai, R.S., C.L. Will, R. Luhrmann, D. Schumperli, and B. Muller. 2001. Purified U7 snRNPs lack the Sm proteins D1 and D2 but contain Lsm10, a new 14 kDa Sm D1-like protein. *EMBO J.* 20:5470–5479.
- Pillai, R.S., M. Grimmmler, G. Meister, C.L. Will, R. Luhrmann, U. Fischer, and D. Schumperli. 2003. Unique Sm core structure of U7 snRNPs: assembly by a specialized SMN complex and the role of a new component, Lsm11, in histone RNA processing. *Genes Dev.* 17:2321–2333.
- Ranish, J.A., E.C. Yi, D.M. Leslie, S.O. Purvine, D.R. Goodlett, J. Eng, and R. Aebersold. 2003. The study of macromolecular complexes by quantitative proteomics. *Nat. Genet.* 33:349–355.
- Rosoll, W., A.K. Kroning, U.M. Ohndorf, C. Steegborn, S. Jablonka, and M. Sendtner. 2002. Specific interaction of Smn, the spinal muscular atrophy determining gene product, with hnRNP-R and gry-rbp/hnRNP-Q: a role for Smn in RNA processing in motor axons? *Hum. Mol. Genet.* 11:93–105.
- Rothbauer, U., K. Zolghadr, S. Muyldermans, A. Schepers, M.C. Cardoso, and H. Leonhardt. 2008. A versatile nanotrap for biochemical and functional studies with fluorescent fusion proteins. *Mol. Cell. Proteomics.* 7:282–289 Epub 2007 Oct 21.
- Selbach, M., and M. Mann. 2006. Protein interaction screening by quantitative immunoprecipitation combined with knockdown (QUICK). *Nat. Methods.* 3:981–983 Epub 2006 Oct 29.
- Sleeman, J.E., L. Trinkle-Mulcahy, A.R. Prescott, S.C. Ogg, and A.I. Lamond. 2003. Cajal body proteins SMN and Coilin show differential dynamic behaviour in vivo. *J. Cell Sci.* 116:2039–2050 Epub 2003 Apr 1.
- Tackett, A.J., J.A. DeGrasse, M.D. Sekedat, M. Oeffinger, M.P. Rout, and B.T. Chait. 2005. I-DIRT, a general method for distinguishing between specific and nonspecific protein interactions. *J. Proteome Res.* 4:1752–1756.

- Trinkle-Mulcahy, L., J. Andersen, Y.W. Lam, G. Moorhead, M. Mann, and A.I. Lamond. 2006. Repo-Man recruits PP1 gamma to chromatin and is essential for cell viability. *J. Cell Biol.* 172:679–692 Epub 2006 Feb 21.
- Vermeulen, M., N.C. Hubner, and M. Mann. 2008. High confidence determination of specific protein-protein interactions using quantitative mass spectrometry. *Curr. Opin. Biotechnol.* 19:331–337 Epub 2008 Jul 11.
- Wang, X., and L. Huang. 2008. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics.* 7:46–57 Epub 2007 Oct 12.
- Yong, J., L. Pellizzoni, and G. Dreyfuss. 2002. Sequence-specific interaction of U1 snRNA with the SMN complex. *EMBO J.* 21:1188–1196.